DOCUMENT RESUME

ED 151 400 TM 006 862

AUTHOR Irvine, David J.; And Others
TITLE Hypothesis Testing in Evaluating a Prekindergarten Program: Illustrative Procedures. Technical Note No. 1.
INSTITUTION New York State Education Dept., Albany. Div. of Research.
PUB DATE 30 Jul 76
NOTE 9p.; For related documents, see TM 006 634, 861, 863, and 903

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Age Differences; Analysis of Covariance; Cognitive Development; Cognitive Tests; Disadvantaged Youth; Early Childhood Education; *Hypothesis Testing; Individual Development; *Preschool Evaluation; *Preschool Programs; Pretesting; *Program Effectiveness; *Research Design; *Statistical Analysis
IDENTIFIERS New York; NY Experimental Prekindergarten Program

ABSTRACT

Identifying the effects of the New York State Experimental Prekindergarten Program on the development of children requires that a broad array of data be analyzed. In studying the effects of the program, the evaluation unit is interested in determining whether there is an overall effect, and whether children of different ages and different levels of initial development are affected in the same way. An illustration is given to clarify how the process of hypotheses testing reduces the number of assumptions one has to accept in drawing conclusions about the program effects. The analysis of covariance procedure illustrated involves 18 possible hypotheses; however, the branching nature of the analysis results in the testing of no more than six hypotheses. The result obtained in testing a given hypothesis leads to another hypothesis while simulataneously precluding other hypotheses. Thus, efficiency of design is achieved by considering a large number of hypotheses while testing only a limited number of them. At the same time, the design makes it possible to retain the most general hypothesis which adequately explains the data. (Author/MV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED151400

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

David J. Irvine

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

HYPOTHESIS TESTING IN EVALUATING A PREKINDERGARTEN PROGRAM: ILLUSTRATIVE PROCEDURES

David J. Irvine
Thomas L. Hick
Mary D. Horan

The University of the State of New York
THE STATE EDUCATION DEPARTMENT
Division of Research
Prekindergarten Evaluation Unit
Albany, New York 12234

Technical Note #1

July 30, 1976

TM006 862

HYPOTHESIS TESTING IN EVALUATING A PREKINDERGARTEN PROGRAM: ILLUSTRATIVE PROCEDURES

Identifying the effects of the Experimental Prekindergarten Program on the development of children requires that a broad array of data be analyzed.

The simplest analysis would be to compare scores of a group of children on a measure of development at the beginning of their prekindergarten year with scores on the same measure at the end of the year. Obviously, this procedure leaves much to be desired since it does not provide any way of accounting for other possible influences on the children's development. To accept the results of this simple design as an indication of the effects of the prekindergarten program requires that one also accept a number of assumptions, including assumptions that such factors as maturation, family conditions, and general ability of the children do not affect the performance of the children on the measure being used. In fact, this simple design requires one to accept such a vast number of assumptions that they could be summed under one general assumption: Nothing in the world outside the prekindergarten program affects the development of children on the measure in question. For purposes of making decisions about a program, this degree of uncertainty about the possible influences on children's development is unsatisfactory. It may also be unnecessary.

While it is not possible to eliminate all assumptions, a large number of them can be brought out of the realm of speculation. The data which can be obtained and the nature of the analytical design will determine which assumptions can be tested and which must remain assumptions. The aim of research is to reduce

both the number and the logical importance of the assumptions that remain unexamined, so that the consumer of the research findings has to accept as little as possible "on faith."

Several steps are followed in examining a group of assumptions: They are stated as null hypotheses; data are collected on variables referred to in the hypotheses; and statistical tests are carried out to determine whether the data warrant retaining or rejecting each hypothesis.

An illustration may clarify how the process of testing hypotheses reduces the number of assumptions one has to accept in drawing conclusions about the effects of a program.

In studying the effects of the Experimental Prekindergarten Program, the evaluation unit is interested not only in determining whether there is an overall effect but also whether it affects children of different ages and different levels of initial development in the same way. In other words, it is not necessary to assume that the program affects children of different ages and levels of development in the same way, since these assumptions are testable and could have meaning for the way in which the program is carried out. Nor is it necessary to assume that age and level of development do not interact in some fashion, since that assumption is also testable and could have implications for the program.

Testing these hypotheses is an integral part of examining Question 1 in the Evaluation Design* for the Experimental Prekindergarten Program. Question 1 asks: Do Prekindergarten programs have a measurable effect upon children's cognitive and noncognitive development between the time they enter and leave the program?

*Evaluation Design: New York State Experimental Prekindergarten Program. Albany, New York: The State Education Department, Bureau of Child Development and Parent Education, Bureau of School Programs Evaluation. September 3, 1975.

The following is one of many possible analytic procedures for answering the question. It is presented here to illustrate the kind of analyses which can be used to study this and other questions about the Prekindergarten Program. In order to keep this example relatively simple, it is assumed that any effects of control variables are linear.

As will be seen, this analysis involves 18 possible hypotheses; however the branching nature of the analysis results in the testing of no more than six hypotheses. The result obtained in testing a given hypothesis leads to another hypothesis while simultaneously precluding other hypotheses. Thus, efficiency of design is achieved by considering a large number of hypotheses while being required to test only a limited number of them. At the same time, the design makes it possible to retain the most general hypothesis which adequately explains the data.

The example described below follows analysis of covariance procedures. The variables used are: Walker posttest (criterion); Walker pretest and children's ages at entry to the program (control variables); and Prekindergarten experience (treatment).

The hypotheses are presented in Table 1.

At the outset, the example seeks to establish the usefulness of the control variables (Hypotheses 1 and 2). Control variables are dropped if they do not contribute to an understanding of a Prekindergarten effect. This is the case at Hypothesis 3 which is answered by the familiar t-test. If the t-test is significant, the hypothesis is rejected and one may conclude that there is a Prekindergarten effect.

Table 1

Illustrative Analysis of Covariance
with two Control Variables (Walker and Age)

Hyp. 1: No effect on the Walker pretest within age for either the experimental or the control group.
Retain go to 2
Reject go to 6

Hyp. 2: No effect of age in either the experimental group or the control group.
Retain go to 3
Reject go to 4

Hyp. 3: No effect of Prekindergarten Program (experimental mean compared to control mean).
Retain - Stop
Reject - Stop

Hyp. 4: No Prekindergarten by age interaction.
Retain go to 5
Reject - Stop

Hyp. 5: No Prekindergarten effect with control for age.
Retain - Stop
Reject - Stop

Hyp. 6: No Walker pretest by age interaction.
Retain go to 7
Reject go to 14

Hyp. 7: No Prekindergarten effect by Walker pretest interaction.
Retain go to 8
Reject go to 12

Hyp. 8: No effect of age with control for Walker pretest.
Retain go to 9
Reject go to 10

Hyp. 9: No Prekindergarten effect with control for Walker pretest.
Retain - Stop
Reject - Stop

Table 1 (continued)

Hyp. 10: No Prekindergarten effect by age interaction in the presence of a homogeneous Walker pretest effect.
Retain go to 11
Reject - Stop

Hyp. 11: No Prekindergarten effect at all levels of age and Walker pretest.
Retain - Stop
Reject - Stop

Hyp. 12: No age effect in the presence of a treatment pretest interaction.
Retain - Stop
Reject go to 13

Hyp. 13: No Prekindergarten effect by age interaction in the presence of treatment by Walker pretest interaction.
Retain - Stop
Reject - Stop

Hyp. 14: No interaction of the Prekindergarten effect and age and Walker pretest.
Retain go to 15
Reject - Stop

Hyp. 15: No Prekindergarten effect by Walker pretest interaction.
Retain go to 16
Reject go to 18

Hyp. 16: No Prekindergarten effect by age interaction in the presence of Walker pretest by age interaction.
Retain go to 17
Reject - Stop

Hyp. 17: No Prekindergarten effect at all levels of age and Walker pretest in the presence of Walker pretest by age interaction.
Retain - Stop
Reject - Stop

Hyp. 18: No Prekindergarten effect by age interaction in the presence of pretest interactions.
Retain - Stop
Reject - Stop

The case where one control variable (age) has a simple effect and the other (Walker pretest) has none at all, is given at Hypothesis 5. Analysis for this hypothesis is again a familiar analysis of covariance with one control variable. Since this analysis (Hypothesis 5) was produced only after testing for a Walker pretest effect, the results of this analysis, and the possible conclusion that Prekindergarten has an effect, hold over the entire range of the rejected variable (Walker Pretest). The corresponding hypothesis, where Walker pretest is retained as a control variable while the age variable was eliminated, is given in Hypothesis 9.

Complex effects of Prekindergarten are located in interaction hypotheses. The most complex, of course, is the triple interaction, expressed in Hypothesis 14. A STOP signal is given after some of these complex effects even though additional hypotheses could be tested.

In general, a STOP signal also designates a point at which conclusions may be drawn. Some of these have been mentioned above in relation to Hypotheses 3 and 5. A significant finding at Hypothesis 14, which would cause rejection of the hypothesis, would lead to conclusions that the Prekindergarten effect was different for children of different ages and of different score levels on the pretest. It might be possible to plot this to show that Prekindergarten was most effective for children entering the program between certain ages and scoring between certain raw score points on the Walker pretest. The data, of course, would determine the precise nature of such a conclusion.

Thus, the results of the analyses may help answer questions of differing breadth of impact from "Should the program be continued?" through "How can the guidelines of the program be changed to make the program more effective?" to "What specific practices are most effective with which children?"

Why is a complex analysis required to answer a rather straight-forward question such as Question 1?

Attempting to answer the general question of effects of the Prekindergarten Program without considering many other possible influences on children's development could lead to an incorrect answer. Differences resulting from other influences could be attributed to the Prekindergarten Program, with the result that it is incorrectly concluded that the program has been effective. Or other influences could obscure the real effects of the Prekindergarten Program so that it is incorrectly concluded that the program has not been effective; this is likely to happen if, for example, the program is more effective for children of a given age or a given level of initial performance than it is for children of other ages or performance levels.

But assuming that the general question is answered correctly, it is important to know how specific factors affect different children under different circumstances. It is possible for a program to be generally ineffective but for specific aspects of the program to be effective; or for the program to be effective with some children but not with others. Even with a program which has been found to be effective overall, it is unlikely that all aspects of the program are equally effective with all children. Efficiency can be improved by an analysis of program variations in relation to children's characteristics.

A great variety of data is being collected on the Experimental Prekindergarten Program to make it possible to carry out analyses similar to that described above. Throughout the analyses, an effort will be made to obtain meaningful answers to the research questions while doing justice to the complexities of the data and the phenomena they represent.